

乳酸菌基因组学研究进展

张文羿¹, 孟和^{2*}, 张和平^{1*}

(¹ 乳品生物技术与工程教育部重点实验室, 内蒙古农业大学食品科学与工程学院, 呼和浩特 010018)

(² 上海交通大学农业与生物学院, 上海 200240)

摘要: 伴随着高通量测序技术的快速发展和测序成本的降低, 越来越多的微生物基因组全序列测定得以实现, 从基因组学的层面了解乳酸菌的遗传结构和组成, 进而分析和掌握其生物学功能已经逐渐成为可能。迄今为止已经有 22 株乳酸菌的基因组完成测序并公开发表, 还有至少 12 株测序工作仍在进行中。本文在分析相关文献和生物信息数据基础上, 从乳酸菌基因组特点、代谢多样性、进化及共线性四个方面对乳酸菌基因组学研究进展进行了总结, 旨在为乳酸菌研究和应用提供参考。

关键词: 乳酸菌; 基因组; 代谢; 进化; 共线性

中图分类号: Q93 文献标识码: A 文章编号: 0001-6209 (2008) 09-1270-06

乳酸菌(Lactic acid bacteria, LAB)是一类革兰氏染色呈阳性, 发酵己糖以乳酸为主要代谢终产物的兼性厌氧细菌的总称, 其中包括致病性链球菌和非致病性乳酸细菌, 根据代谢产物的种类还可将乳酸菌分为同型发酵乳酸菌和异型发酵乳酸菌。目前应用于工业生产的乳酸菌菌种主要是: 乳球菌(*Lactococcus*)、乳杆菌(*Lactobacillus*)、肠膜明串珠菌(*Leuconostoc*)、片球菌(*Pediococcus*)、粪肠球菌(*Enterococcus faecalis*)、酒类酒球菌(*Oenococcus oeni*)和嗜热链球菌(*Streptococcus thermophilus*)。乳酸菌可以栖居于人和动物肠道及各种器官中, 能够调节肠道微生态平衡, 具有抗肿瘤、降胆固醇、延缓衰老等功效^[1]。一些具有益生特性的乳酸菌作为活菌制剂可以激活体内机体免疫系统中的巨噬细胞, 从而提高特异和非特异性免疫应答发挥免疫疗效作用^[2]。乳酸菌发酵食品已有悠久的历史, 乳酸发酵已被广泛应用于乳制品, 肉制品, 青贮饲料, 谷物和果蔬等的加工生产^[3,4]。

早期乳酸菌的研究工作多集中于菌种的分类和菌株的筛选, 并通过构建动物模型对部分益生特性进

行验证。随着分子生物学理论和生物技术的发展, 人们开始利用分子生物学方法和手段对乳酸菌分类鉴定、基因功能及其代谢途径进行分析研究。由于这些方法和手段往往以单个基因和途径为目标, 能够分析的数据信息量过少, 很难对基因和代谢网络调控体系做出令人满意的分析和解释。1995年 Fleischman 等采用全基因组鸟枪法测序(whole genome shotgun sequencing)对嗜血流感菌(*Haemophilus influenzae*)进行测定并获得了成功^[5], 此后这种快速而有效测序方法被广泛应用于微生物基因组测定。第一株乳酸菌即乳酸乳球菌(*Lactococcus lactis* subsp. *lactis* IL1403)全基因组测序项目于 1999 年启动并在 2001 完成^[6]。此后陆续开展了其他乳酸菌全基因组测序和分析工作, 这个时期乳酸菌基因组测序主要针对链球菌基因组^[7,8]。真正意义上乳酸菌大规模测序始于 2002 年召开的乳酸菌基因组专题讨论会, 从此乳酸菌研究发生了革命性的转变, 仅 2005-2006 年, 就先后公布了 13 株乳酸菌的全基因组数据, 测序技术的发展极大地推动了整个乳酸菌基因组学的发展进程^[9,10]。

基金项目: 国家自然科学基金(30560097, 30660135, 30760156); 国家“863 计划”(2006AA10Z345, 2007AA10Z353); 教育部新世纪优秀人才支持计划资助(NCET-06-0269)

*通讯作者。张和平, Tel: +86-471-4319940, Fax: +86-471-4300122, E-mail: hepingdd@vip.sina.com; 孟和, Tel: +86-21-34204538, Fax: +86-21-34204107, E-mail: menghe@sjtu.edu.cn

作者简介: 张文羿(1981-), 男, 内蒙古人, 博士研究生, 研究方向为乳品生物技术。E-mail: zhangwenyizi@163.com

收稿日期: 2008-03-10; 修回日期: 2008-04-25

截止 2007 年 11 月共有 22 株乳酸菌基因组相继完成全序列测定并向国际公共数据库递交,随之而来的是有关乳酸菌蛋白质组(proteome)、转录组(transcriptome)及代谢物组(metabolome)相关术语和数据在文献中的大量显现^[9, 11~13],这些新思路和新方法也为人们深入全面了解和认识乳酸菌揭开了新篇章。很显然,研究单个或少数几个基因和蛋白质的方法,正在被基因组学及其系统生物学方法所取代。生物信息学的兴起和迅猛发展也为乳酸菌全基因组测序和研究大开方便之门,乳酸菌基因组研究也被赋予了更广泛的含义,成为乳酸菌研究的重点和热点。日益庞大的乳酸菌基因组数据库无疑将有助于理解乳酸菌的进化及其同自然生态位之间的关系。

1 乳酸菌基因组的全基因组测序及其基本特征

乳酸菌基因组相对较小且重复序列较少,实践证明鸟枪法测序策略可以方便、快捷的完成其基因组测序任务。由美国能源部基因研究所(Joint Genome Institute, JGI)与乳酸菌基因组协会(Lactic Acid Bacteria Genome Consortium, LABGC)合作完成的 9 株乳酸菌,测序过程中几乎采用同样的方法,即首先对较小插入片段的文库(3 kb)和较大片段(40 kb)的粘粒文库进行测序,使测序反应数的覆盖率分别达到 10 × 和 5 ×,然后再用所得数据进行组装和拼接,后期完成阶段采用 PCR 引物步移法(primer-walking)进行序列缺口的填补。与传统的测序法相比,这种测序方法省去了许多中间步骤,能够快速的对基因组进行测序。由于计算机能力和拼接软件功能的不断提高,用这种方法对乳酸菌基因组进行测序也越来越普遍。

迄今为止,已发表的乳酸菌基因组包括:嗜酸乳杆菌(*Lactobacillus acidophilus* NCFM),植物乳杆菌(*Lactobacillus plantarum* WCFS1),短乳杆菌(*Lactobacillus brevis* ATCC367),约氏乳杆菌(*Lactobacillus johnsonii* NCC533),干酪乳杆菌(*Lactobacillus casei* ATCC334),沙克乳杆菌(*Lactobacillus sakei* 23K),格氏乳杆菌(*Lactobacillus gasseri* ATCC33323),唾液乳杆菌(*Lactobacillus salivarius* subsp.*salivarius* UCC118)等 22 株乳酸菌,大部分由美国和欧洲等发达国家的实验室完成,包括一些知名食品企业参与资助的测序项目。在我国,由内蒙古农业大学“乳品生物技术与工程”教育部重点

实验室和中国科学院北京基因组研究所合作的干酪乳杆菌(*Lactobacillus casei* Zhang)基因组测序工作已进入后期完成阶段,该菌也是国内较为系统开发的益生菌菌株之一。

不同种类的乳酸菌基因组序列的长度在 1.8~2.9 Mb 之间,只有干酪乳杆菌 ATCC334 和植物乳杆菌 WCFS1 的基因组相对较大,分别为 2.95 Mb 和 3.35 Mb。大多数乳酸菌基因组 G+C 含量都在 50% 以下,唾液乳杆菌 UCC118 的 G+C 含量最低只有 33%,而双歧杆菌(*Bifidobacterium*)的 G+C 含量则高达 60.1%。大多数已完成测序的乳酸菌细胞中都含有不同数量的质粒,质粒编码基因在基因组中占 0%~4.8%不等,质粒大小从 2~242 kb 不等^[13,14]。其中,明串珠菌 ATCC8293 和部分乳杆菌细胞中有质粒出现,德氏乳杆菌和大多数嗜热链球菌细胞中均没有质粒,乳酸乳球菌 SK11 细胞中有 5 个不同种类的质粒。此外,乳酸菌基因组中都含有一定比例的插入序列和原嗜菌体序列^[15]。

2 乳酸菌基因组代谢的多样性

代谢作用中的化学反应是生物体从周围环境获得能量的主要方式,通过化学诱变和遗传重组等改良技术,结合部分生化数据研究这些反应网络体系的代谢流分布是乳酸菌代谢建模的基本思路。这在早期丙酮酸盐代谢途径建模的研究过程中也被证明是行之有效的,然而在解释复杂的生物合成途径过程中却略显不足,这在一定程度上妨碍了对乳酸菌代谢能力的改造^[16]。为了更好的了解乳酸菌细胞的代谢途径,使乳酸菌生物工程育种做到有的放矢,引入更多的参数建立更为强大的动力学模型也成为不时之需。已获得的基因组数据,涵盖了用于工业生产的乳酸菌,反应了不同菌属乳酸菌代谢的多样性,也为重建乳酸菌代谢途径提供了良好的遗传学基础。

近期报道指出不同的乳酸菌所含的酶系不同,代谢途径及形成的最终产物也不同。通过对一些乳酸菌进行基因组测序,也发现了一些未预计到的代谢途径。乳酸菌主要通过厌氧糖酵解途径(Embden-Meyrhoof-Parman Pathway)来获取能量,但研究人员也发现乳酸乳球菌 IL1403 基因组中还编码有氧呼吸的酶类,这也表明可能有其它产能的代谢途径^[6,17~20]。嗜酸乳杆菌、约氏乳杆菌,保加利亚乳杆菌基因组中都不存在丙酮酸盐复合物体系,也不编码

可以在厌氧条件下直接能够以丙酮酸盐为底物产生乙酰辅酶 A 的酶类,植物乳杆菌拥有更完善的丙酮酸盐代谢系统(pyruvate-dissipating)^[21-24]。然而,乳酸乳球菌也编码有冗余的乳酸盐脱氢酶的基因,并且乳酸乳球菌和植物乳杆菌基因组中都缺乏完整的柠檬盐循环^[13]。在鉴定为亲缘关系较远的乳酸菌中,植物乳杆菌、乳酸乳球菌和嗜热链球菌在不同碳源的生长环境中,却有着类似的碳代谢蛋白(carbon catabolite protein A, ccpA) 响应机制^[25]。

全基因组分析显示,植物乳杆菌基因组编码有大量的磷酸转移酶系统(phosphotransferase system, PTS)以及其它的糖转运系统相关酶类的基因,这些成员中有 30 个转运蛋白被认为与摄取和利用碳源相关^[26]。在嗜酸菌群中,包括约氏乳杆菌和格氏乳杆菌的基因组,以及球菌类群中的乳酸乳球菌在内也存在类似比例编码 PTS 的基因,仅有为数不多其它类型的转运系统^[27]。对于嗜酸乳杆菌而言,乳糖和半乳糖是由半乳糖苷酶通透酶转运到胞内,但基因组中仅发现三个对麦芽糖和低聚果糖有潜在转运作用的基因^[22]。这也符合保加利亚乳杆菌、植物乳杆菌等同型发酵乳酸菌在进入酵解之前,单糖一般都要进行磷酸化,并且由磷酸转移酶系统使单糖转移进入细胞的事实。而一些异型发酵乳酸菌,如肠膜明串珠菌、短乳杆菌和嗜热链球菌等单糖则是在乳酸透性酶运输系统的作用下进入细胞进行水解和磷酸化^[28]。

从基因组序列得出乳酸菌代谢另一特点是,不同种属乳酸菌氨基酸合成途径有不同程度的缺失和已知转运体同源物的表观功能的缺失。约氏乳杆菌、嗜酸乳杆菌和格氏乳杆菌基因组中编码的氨基酸生物合成途径较少,它们都缺乏维生素和嘌呤核苷酸生物合成必需酶,编码参与糖代谢,氨基酸、核酸和脂肪酸生物合成的蛋白数量较少^[8]。但基因组中大量编码的肽酶,氨基酸通透酶和寡肽转运系统基因也许可以在营养丰富环境中行使氨基酸摄取的功能,补偿氨基酸合成途径的缺失。与此相反,植物乳杆菌编码相对完整的生物合成途径,除亮氨酸,异亮氨酸,缬氨酸之外,可自身合成几乎所有的氨基酸^[29]。类似的,乳酸乳球菌也呈现出很强的合成氨基酸能力,拥有合成 20 种氨基酸的遗传学潜力,可利用约 20 种化合物作为底物碳源,涉及到 83 个调节因子,这也使它能够能够在缺少氨基酸的环境中生存^[6,30]。

比较栖息生境紧密相关乳酸菌代谢内容的差别,

比如说,在工业上常常被用于混合发酵牛乳的保加利亚乳杆菌和嗜热链球菌,也许有助于我们理解它们在发酵过程中的互惠共生作用。保加利亚乳杆菌几乎不具备生物合成氨基酸的能力,而嗜热链球菌除组氨酸之外几乎保留了所有与氨基酸合成相关的酶。这也许可以解释为保加利亚乳杆菌较强的蛋白水解活性为嗜热链球菌的生长提供必要的短肽和氨基酸的生化现象^[24,31]。并且和嗜酸菌群的成员一样,都编码有鸟氨酸脱羧酶可以催化鸟氨酸的脱羧反应。因为这并不是保加利亚乳杆菌生长过程中的必须基因,由此推测,在乳制品发酵过程中鸟氨酸由嗜热链球菌来提供,并且嗜热链球菌基因组中有两个 ABC 家族的转运蛋白底物为丁二胺,从而认为二者在同处一个环境时,嗜热链球菌为保加利亚乳杆菌生长提供了需要的生长因子促进共生。

3 种系发生分析和乳酸菌基因组的进化

现存乳酸菌的进化关系一直是乳酸菌生物学研究的重点,也是多少年以来的难点。在研究早期,乳酸菌种属间在系统发育上的关系主要通过比较菌落形态特征和糖类发酵反应的不同来加以区分。随着分子生物学技术的发展,在分类上,将表型鉴定结果和遗传物质结构相结合来说明属种在系统发育上的关系,并通过定量的分析进一步澄清了由传统分类学确立的不同乳酸菌种属之间的进化关系^[32]。但是事实上,仅以单个基因为研究对象得到的进化关系并不能真实的反映乳酸菌在自然界各种生境中的进化关系^[33]。不同标记构建出的种系发生树也不尽相同,有人将其归因于水平基因转移,因为水平转移的基因将会模糊物种间的进化关系^[34]。因此,将多个基因组合起来进行系统发育学的分析,可以降低噪声,得到的结论可能更为客观。

近几年来,发表的乳酸菌基因组全序列逐年增加,为我们提供了更多的核酸序列变异的资料,从基因组水平来探索种系发生史已成为可能,这种研究方法也代表了当今乳酸菌进化领域的一种趋势。有了可靠的系统发生树,就可以重建较好的不同乳酸菌种属间的进化关系。根据这些思路,运用多种方法如最大似然法、及临近原则构建的系统进化树拓扑结构相同,并且与基于核糖体 RNA 构建的系统进化树高度一致。在这些系统发育树中,很明显乳酸菌中至少可分为 3 个主要类群,链球菌属和乳球菌属为第一类群,

干酪乳杆菌和嗜酸菌群成员构成第二类群, 片球菌属和明串珠菌属都属于同一个进化分支, 与短乳杆菌以及植物乳杆菌共同组成另一类群^[15]。由于复杂的进化过程和历史, 对直向同源基因的鉴别是构建正确系统进化树的关键, 这对功能和进化研究也是必不可少的^[35]。

Makarova 及其同事将构建的数据集—乳酸菌蛋白质直向同源群的聚类(Lactobacillales-specific clusters of orthologous, LaCOGs)和种系发生树结合起来重建了乳酸菌的进化史^[36]。该研究小组的分析表明, 可适应性进化机理在乳酸菌的基因组主体结构形成过程中发挥了极其重要的作用并促成了“益生菌”种的形成。在营养丰富环境的适应过程中, 合成途径中的相关酶类不同程度的缺失, 最终使乳酸菌出现了营养缺陷的表型特征。在不同的谱系中, 如包括德氏乳杆菌, 格氏乳杆菌和约氏乳杆菌的嗜酸菌群, 基因组容量较大的植物乳杆菌和干酪乳杆菌基因组中, 都发生过基因的大面积缺失和失活突变。尤其是在球菌属和乳球菌属中, 象 *crcB*, *mreB*, *mreC* 和 *minD* 这样公认的在细菌中是非常保守的与细胞分裂有关的基因也存在丢失的现象。这些结果提示基因缺失在乳酸菌进化过程中起到了很重要的作用。其它对于乳酸菌基因组的研究, 特别是在嗜热链球菌中, 也可见证这种机制的存在^[37]。

比较基因组学得出的另一个重要的结论, 即基因水平转移(horizontal gene transfer, HGT)在乳酸菌基因组进化过程中也扮演了重要的角色。基因组分析显示, 基因水平转移在乳酸菌基因组中是个普遍的进化事件。大量证据显示乳球菌细胞中已获得了在牛奶中生长和竞争的特定功能的质粒 DNA 片段, 这些功能包括乳糖代谢、水解蛋白、产生表面多糖、噬菌体抗性。嗜酸乳杆菌合成胞外多糖的能力, 植物乳杆菌中编码糖转运、代谢和表面多糖合成的基因, 也都被证明是在适应各种环境的过程中通过基因水平转移获得的^[12,22]。此外, 在其它乳酸菌基因组中, 例如, 乳酸乳球菌 MG1363 和乳酸乳球菌 IL1403 的染色体上, 通过转化、噬菌体介导的转移和接合转移而获得的外源 DNA 也占有一定比重, 分别为编码基因的 7.1% 和 8.9%^[6,30]。这些可移动功能基因的转入, 在赋予了受体乳酸菌新的代谢特征的同时, 也极大地推动了乳酸菌基因组的多样化进程。

4 乳酸菌基因组间的共线性

基因组共线性(colinearity)包括基因组的相似性

和基因排列序列的一致性。随着现代生物信息学的快速发展, 全基因组比较逐渐变成基因组分析的强大工具, 基因组共线性的研究也演变为基因组研究的重要策略之一。应用一些快速作图比对程序, 对两个完整的基因组进行比较, 可以揭示生物体之间的大规模相似性, 也为乳酸菌基因组进化史的研究提供了新的视角。

比较遗传学研究表明, 在遗传图谱水平, 许多细菌基因组所含基因及基因的顺序均高度保守, 但在微观水平上, 共线性水平的保守性不高, 只有在进化距离非常近的物种间才保持良好的共线性。乳酸菌全基因组比较的结果表明, 近缘物种菌株间存在大量的相同序列。在乳杆菌属中, 约氏乳杆菌和嗜酸乳杆菌在基因组构成和基因组长度之间表现出高度的保守性^[38]。相比之下, 约氏乳杆菌和格氏乳杆菌间具有更高层次的共线性, 仅在两个基因组的复制终止点附近观察到两个大的区段的倒位现象^[24]。植物乳杆菌与约氏乳杆菌、嗜酸乳杆菌、格氏乳杆菌、清酒乳杆菌、唾液乳杆菌之间缺乏共线性。与亲缘关系稍远的球菌属家族成员相比, 约氏乳杆菌与短乳杆菌和片球菌基因组之间的相似程度更高, 而与此同时, 酒类酒球菌和明串珠菌属与约氏乳杆菌间几乎不存在共线性^[15]。与核酸水平比较的结果略有不同, 在蛋白水平上, 嗜酸乳杆菌、约氏乳杆菌和格式乳杆菌基因组间存在广泛的共线性。植物乳杆菌与约氏乳杆菌、嗜酸乳杆菌、格氏乳杆菌、清酒乳杆菌、唾液乳杆菌之间都有一定的共线性。正如大规模研究显示的结果那样, 对乳酸菌菌株的比较都显示出 X 形状的排列样式, 在复制原点处均可观察到 X 形状的排列区。即使在亲缘关系较远的物种之间, 这种样式仍依稀可辨, 这与核酸水平上的比较分析结果形成了鲜明的对比^[39]。值得注意的是, 在对同一种属中的不同乳酸菌菌株基因组进行比较时, 尽管不同菌株间在宏观上都具有高度保守的基因共线性, 但也往往伴有局部的插入、倒位等现象, 使微观同线性只保持在一定的范围内^[15,30,37]。

5 存在问题和展望

我国乳酸菌的研究起步较晚, 在益生菌的开发应用方面原创性成果较少。目前围绕乳酸菌基因组及其功能基因组学研究几乎还是个空白, 并不能满足国内乳品工业发展日益增长的需求。我们应该将乳酸菌基因组图谱与生物学功能紧密联系起来, 充分利用基因

组提供的生物信息, 深入发掘国内丰富的乳酸菌资源, 参与国际竞争, 这对我国乳品工业的长远发展也有着重要的战略意义。

随着测序技术不断革新发展和成本的降低, 将极大地丰富乳酸菌基因组数据库资源。利用这些完整的乳酸菌基因组数据, 结合双向凝胶电泳、微阵列和生物信息学技术手段和分析方法, 可以帮助我们全面、系统地了解乳酸菌基因组的结构和组成, 为从系统生物学角度分析乳酸菌的生物学功能及形成机制, 进而为成功改造和选育优良益生乳酸菌奠定基础, 最终使它为人类健康事业做出更大贡献。

参 考 文 献

- [1] Vaughan EE, de Vries MC, Zoetendal EG, *et al.* The intestinal LABs. *Antonie Van Leeuwenhoek*, 2002, 82: 341–352.
- [2] Canche-Pool EB, Cortez-Gomez R, Flores-Mejia R, *et al.* Probiotics and autoimmunity: an evolutionary perspective. *Med Hypotheses*, 2007, 70(3): 657–660.
- [3] McKay LL, Baldwin KA. Applications for biotechnology: present and future improvements in lactic acid bacteria. *FEMS Microbiol Rev*, 1990, 87(1–2): 3–14.
- [4] van Belkum A, Nieuwenhuis EE. Life in commercial probiotics. *FEMS Immunol Med Microbiol*, 2007, 50(3): 281–283.
- [5] Fleischmann RD, Adams MD, White O, *et al.* Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*, 1995, 269(5223): 496–512.
- [6] Bolotin A, Mauger S, Malarme K, *et al.* Low-redundancy sequencing of the entire *Lactococcus lactis* IL1403 genome. *Antonie Van Leeuwenhoek*, 1999, 76(1): 27–76.
- [7] Yother J, Trieu-Cuot P, Klaenhammer TR, *et al.* Genetics of streptococci, lactococci, and enterococci: review of the sixth international conference. *J Bacteriol*, 2002, 184(22): 6085–6092.
- [8] Klaenhammer TR, Peril AA, Barrangou R. Genomic perspectives on probiotic lactic acid bacteria. *Biosci Microflora*, 2005, 24(2): 31–33.
- [9] Liu M, van Enckevort FHJ, Siezen RJ. Genome update: lactic acid bacteria genome sequencing is booming. *Microbiology*, 2005, 151(12): 3811–3814.
- [10] Claesson MJ, van Sinderen D, Toole PWO. The genus *Lactobacillus*-a genomic basis for understanding its diversity. *FEMS Microbiol Lett*, 2007, 269(1): 22–28.
- [11] Sonnenburg JL, Chen CTL, Gordon JI. Genomic and metabolic studies of the impact of probiotics on a model gut symbiont and host. *PLoS Biol*, 2006, 4(12): 2213–2226.
- [12] Koistinen KM, Plumed-Ferrer C, Lehesranta SJ, *et al.* Comparison of growth-phase-dependent cytosolic proteomes of two *Lactobacillus plantarum* strains used in food and feed fermentations. *FEMS Microbiol Lett*, 2007, 273(1): 12–21.
- [13] Pfeiler EA, Klaenhammer TR. The genomics of lactic acid bacteria. *Trends Microbiol*, 2007, 15(12): 546–553.
- [14] Klaenhammer T, Altermann E, Arigoni F, *et al.* Discovering lactic acid bacteria by genomics. *Antonie Van Leeuwenhoek*, 2002, 82(1–4): 29–58.
- [15] Makarova K, Sleasarev A, Wolf Y, *et al.* Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci U S A*, 2006, 103(42): 15611–15616.
- [16] Smid EJ, van Enckevort FJH, Wegkamp A, *et al.* Metabolic models for rational improvement of lactic acid bacteria as cell factories. *J Appl Microbiol*, 2005, 98(6): 1326–1331.
- [17] Kok J, Buist G, Zomer A L, *et al.* Comparative and functional genomics of lactococci. *FEMS Microbiol Rev*, 2005, 29(3): 411–433.
- [18] Schell MA, Karmirantzou M, Snel B, *et al.* The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. *Proc Natl Acad Sci U S A*, 2002, 99(22): 14422–14427.
- [19] Claesson MJ, Li Y, Leahy S, Canchaya C, *et al.* Multireplicon genome architecture of *Lactobacillus salivarius*. *Proc Natl Acad Sci U S A*, 2006, 103(17): 6718–6723.
- [20] Chaillou S, Champomier-Verges MC, Cornet M, *et al.* The complete genome sequence of the meat-borne lactic acid bacterium *Lactobacillus sakei* 23K. *Nat Biotechnol*, 2005, 23(12): 1527–1533.
- [21] Coenye T, Vandamme P. Extracting phylogenetic information from whole-genome sequencing projects: the lactic acid bacteria as a test case. *Microbiology*, 2003, 149(Pt12): 3507–3517.
- [22] Altermann E, Russell WM, Azcarate-Peril MA, *et al.* 2004. Complete genome sequence of the probiotic lactic acid bacterium *Lactobacillus acidophilus* NCFM. *Proc Natl Acad Sci USA*, 102(11): 3906–3912.
- [23] Pridmore RD, Berger B, Desiere F, *et al.* The genome sequence of the probiotic intestinal bacterium *Lactobacillus johnsonii* NCC533. *Proc Natl Acad Sci USA*, 2003, 101(8): 2512–2517.
- [24] van de Guchte M, Penaud S, Grimaldi C, *et al.* The complete genome sequence of *Lactobacillus bulgaricus* reveals extensive and ongoing reductive evolution. *Proc Natl Acad Sci USA*, 2006, 103(24): 9274–9279.
- [25] van den Bogaard PTC, Kleerebezem M, Kuipers OP, *et al.* Control of lactose transport, beta-galactosidase activity, and glycolysis by CcpA in *Streptococcus thermophilus*: evidence for carbon catabolite repression by a non-phosphoenol-pyruvate dependent phosphotransferase system. *J Bacteriol*, 2000, 182: 5982–5989.
- [26] Kleerebezem M, Boekhorst J, van Kranenburg R, *et al.* Complete genome sequence of *Lactobacillus plantarum* WCFS1. *Proc Natl Acad Sci USA*, 2003, 100(4): 1990–1995.
- [27] Klaenhammer TR, Barrangou R, Buck BL, *et al.* Genomic features of lactic acid bacteria effecting bioprocessing and health. *FEMS Microbiol Rev*, 2005, 29(3): 393–409.

- [28] 王海燕, 刘铭, 王化军, 等. 乳酸生产中的微生物代谢工程. 过程工程学报 (*The Chinese Journal of Process Engineering*), 2006, 5(3): 512–516.
- [29] Boekhorst J, Siezen RJ, Zwahlen MC, *et al.* The complete genome of *Lactobacillus plantarum* and *Lactobacillus johnsonii* reveal extensive differences in chromosome. *Microbiology*, 2004, 150(Pt 11): 3601–3611.
- [30] Wegmann U, Connell-Motherway MO, Zomer A, *et al.* Complete genome sequence of the prototype lactic acid bacterium *Lactococcus lactis* subsp. *cremoris* MG1363. *J Bacteriol*, 189(8): 3256–3270.
- [31] Davidson BE, Kordias N, Dobos M, *et al.* Genomic organization of lactic acid bacteria. *Antonie van Leeuwenhoek*, 1996, 70(2–4): 161–183.
- [32] Stiles ME, Holzapfel WH. Lactic acid bacteria of foods and their current taxonomy. *Int J Food Microbiol*, 1997, 36(1): 1–29.
- [33] London J, Kline K. Aldolase of lactic acid bacteria: a case history in the use of an enzyme as an evolutionary marker. *Bacteriol Rev*, 1973, 37: 453–478.
- [34] Gogarten JP. The early evolution of cellular life. *Trends Ecol Evol*, 1995, 10(4): 147–151.
- [35] Tatusov RL, Natale DA, Garkavtsev LV, *et al.* The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res*, 2001, 29(1): 22–28.
- [36] Makarova KS, Koonin EV. Evolutionary genomics of lactic acid bacteria. *J Bacteriol*, 2007, 189(4): 1199–1208.
- [37] Bolotin A, Quinquis B, Renault P, *et al.* Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nat Biotechnol*, 2004, 22(12): 1554–1558.
- [38] Canchaya C, Claesson MJ, Fitzgerald GF, *et al.* Diversity of the genus *Lactobacillus* revealed by comparative genomics of five species. *Microbiology*, 2006, 152, 3185–3196.
- [39] Berger B, Pridmore RD, Barretto C, *et al.* Similarity and differences in the *Lactobacillus acidophilus* group identified by polyphasic analysis and comparative genomics. *J Bacteriol*, 2007, 189(4): 1311–1321.

Progress on the genomics of lactic acid bacteria—A review

Wenyi Zhang¹, He Meng^{2*}, Heping Zhang^{1*}

¹Key laboratory of Dairy Biotechnology and Engineering, Ministry of Education, Department of Food Science and Engineering, Inner Mongolia Agricultural University, Huhhot 010018, China

²School of Agriculture and biology Shanghai Jiaotong University, Shanghai 200040, China

Abstract: With the advances and cost reducing of DNA-sequencing, genome determination for many microorganisms becomes available. For the research of lactic acid bacteria (LAB), this progress will not only facilitate to understand their nature of hereditary constitution, but also gradually leads to the possibility for analysis and further control of their biological functions. At present, the whole genome for 22 strains of LAB have been completed and published, at least another 12 strains of LAB are under sequencing. Based on the published data, we described the research progression of LAB genomics from four aspects, including general features, diversity of metabolism, evolution and colinearity.

Keywords: lactic acid bacteria (LAB); genome; metabolism; evolution; colinearity

Supported by the National Nature Science Foundation of China (30560097,30660135,30760156), the National Programs for High Technology Research and Development of China (2006AA10Z345, 2007AA10Z353) and the New Century Excellent Talent Planning of Education Ministry of China (NCET-06-0269)

*Corresponding author. Heping Zhang, Tel: +86-471-4319940, Fax: +86-471-4300122, E-mail:hepingdd@vip.sina.com; He Meng, Tel: +86-21-34204538, Fax: +86-21-34204107, E-mail: menghe@sjtu.edu.cn

Received: 10 March 2008/ Revised: 25 April 2008

《微生物学报》答作者问——关于版权

问: 在与贵刊签订版权协议时, 我只想转让印刷版, 不想转让网络版, 是否可以?

答: 本刊不仅出版印刷版, 同时出版光盘版、网络版。在出版过程中, 一般不大可能将其中的一篇文章撤下来不做网络版, 所以凡是不同意转让光盘版和网络版的作者, 本刊建议您改投其他刊物。请您谅解!